Nombre del alumno: Lesly Merari Utrilla López.

Nombre del profesor: Rosario Gómez Lujano.

Nombre del trabajo: Ensayo.

Materia: Estadística.

Grado: Primer cuatrimestre.

Grupo: psicología escolarizado.

Pichucalco, Chiapas a 04 de diciembre de 2020.

I. Investigar y realizar un ensayo de 5 cuartillas de los siguientes temas: **distribuciones de variable discreta, distribución de variable continua, métodos de muestreos, intervalos de confianza.**

La estadística es una de las herramientas que ayudan a la humanidad en diversos aspectos de la vida.

Se pueden identificar diversos procesos que para la estadística son básicos.

Si nosotros requerimos realizar una encuesta que nos demuestre algún resultado de nuestro interés. Necesitamos de una población y de una muestra, pues sabemos que las poblaciones son muy difíciles de entrevistar o de estudiar.

Para llevar a cabo esto necesitamos de métodos que nos ayuden a determinar que porción de la población es la adecuada.

Contamos con dos tipos de muestreo para nuestra fortuna, que a su vez se dividen en pequeñas técnicas.

1. Muestreo probabilístico:

Es un método de muestreo (muestreo se refiere al estudio o el análisis de grupos pequeños de una población) que utiliza formas de métodos de selección aleatoria. El requisito más importante del muestreo probabilístico es que todos en una población tengan la misma oportunidad de ser seleccionados.

Aleatorio simple. Es el método más sencillo de extracción de una muestra probabilística.

El procedimiento empleado es el siguiente:

1) Se asigna un número a cada individuo de la población.
2) A través de algún medio mecánico (bolas dentro de una bolsa, tablas de numeroso aleatorios, etc.) se eligen tantos sujetos como sea posible y necesario para completar el tamaño de muestra requerido.

Este procedimiento, atractivo por su simpleza, tiene poca o nula utilidad práctica cuando la población que estamos manejando es muy grande.

Aleatorio sistemático.

- El procedimiento exige numerar todos los elementos de la población.
- Extraer un numero aleatorio y a partir de ese único número aleatorio y se empiezan a elegir a los elementos que integran a la muestra y serán los que ocupan los lugares i, i + k, i + 2k, i + 3k, (n-1) k, es decir se toman los individuos de k en k.
- K es el resultado de dividir el tamaño de la población entre el tamaño de la población entre el tamaño de la muestra: k=N/n.

- El numero i que empleamos como punto de partida será un numero al azar entre 1 y k.

Aleatorio estratificado. Simplifican los procesos y suelen reducir el error muestra para un tamaño dado de la muestra. Consiste en considerar categorías típicas diferentes entre sí (estratos) que poseen gran homogeneidad respecto a alguna característica (se puede estratificar, por ejemplo, según la profesión, el municipio de residencia, sexo, estado civil).

Aleatorio por conglomerados. Consiste en seleccionar aleatoriamente un cierto número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y en investigar después todos los elementos pertenecientes a los conglomerados elegidos. Cuando los conglomerados son áreas geográficas suele haber de muestra ´por áreas. La unidad muestral es un grupo de elementos de la población que forman una unidad, a la que llamamos conglomerados.

2. Muestreo no probabilístico.

Por conveniencia. El entrevistador selecciona unidades que conforman las muestras, selecciona a participantes ya que están dispuestos y disponibles para ser estudiados. "la muestra no es representativa de la población".

Según criterio. Supone la selección de determinados encuestados para que participen en el estudio se selecciona a esos individuos porque parecen representar a la población a analizar "suelen emplearse en estudios de prueba de productos".

De cuotas. Consisten en fijar "cuotas" que consisten en un número de individuos que reúnen unas determinadas condiciones. Se selecciona a números concretos de encuestados que representan ciertas características de las que se sepa o se suponga que pueden afectar el tema de estudio.

Bola de nieve. Es muy utilizado cuando pretendemos llegar a una población de interés que es pequeña y muy especializada.

Cuando hablamos de las variables que pueden ser encontradas en nuestro experimento, están las discretas y las continuas.

Las discretas van a hacer todos los valores enteros y estos se pueden contar., los continuos son aquellos valores enteros y admiten intermedios, se pueden medir.

Ejemplos de las variables discretas: número de personas, número de camas, mascotas, etc.

Ejemplo de las variables continuas: estatura de las personas, longitud de una cama, peso de una mascota.

Al momento de que nosotros presentemos nuestros resultados lo podemos hacer por medio de gráficas, tablas o algún otro modelo.

Nos vamos a encontrar con dos tipos de distribuciones según sea el caso.

1. La distribución continua.

Describe las probabilidades de los posibles valores de una variable aleatoria continua. Una variable aleatoria continua es una variable aleatoria con un conjunto de valores posibles (conocido como el rango) que es infinito y no se puede contar.

Las probabilidades de las variables aleatorias continuas se definen como el área por debajo de la curva. Por lo tanto, solo los rangos de valores pueden tener una probabilidad diferente de cero. La probabilidad de que una variable aleatoria continua equivalga a algún valor siempre es cero.

Dentro de las distribuciones de variable continuas más importantes encontramos:

1) DISTRIBUCIÓN X^2. (de Pearson) es una distribución de probabilidad continua con un parámetro k que da a conocer los grados de libertad de la variable aleatoria: Un experimento al cual se aplica la distribución de Bernoulli se conoce como Ensayo de Bernoulli o simplemente ensayo, y la serie de esos experimentos como ensayos repetidos.

La distribución χ^2 se puede aplicar de diversas maneras en inferencia estadística, por ejemplo, en la denominada prueba χ^2 utilizada como prueba de independencia y como prueba de bondad de ajuste y en la estimación de varianzas. Está también involucrada en la problemática de estimar la media de una población normalmente distribuida y en el problema de estimar la pendiente de una recta de regresión lineal, a través de su papel en la distribución t de Student, y participa en todos los problemas de análisis de varianza, por su papel en la distribución F de Snedecor, que es la distribución del cociente de dos variables aleatorias independientes con distribución χ^2.

2) Distribución t de Student. es una distribución de probabilidad que se origina del problema de estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño.

Aparece de manera natural al realizar la prueba t de Student para la determinación de las diferencias entre dos medias muestrales y para la construcción del intervalo de confianza para la diferencia entre las medias de dos poblaciones cuando se desconoce la desviación típica de una población y ésta debe ser estimada a partir de los datos de una muestra.

3) Distribución normal. (Distribución de Gauss o distribución gaussiana) es una de las distribuciones de probabilidad de variable continua que con más frecuencia aparece en fenómenos reales.

La gráfica de su función de densidad tiene una forma acampanada y es simétrica respecto de un determinado parámetro. Esta curva se conoce como campana de Gauss.

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos.

4) Distribución beta. Es una distribución de probabilidad continua con dos parámetros a y b cuya función de densidad para valores $0 < x < 1$ es

5) Distribución F. También se la conoce como distribución F de Snedecor (por George Snedecor) o como distribución F de Fisher-Snedecor. La distribución F aparece seguidamente como la distribución nula de una prueba estadística, particularmente en el análisis de varianza.

2. Distribución discreta

Es aquella que describe la probabilidad de ocurrencia de cada uno de los valores de una variable aleatoria discreta. Una variable aleatoria discreta es una variable aleatoria que tiene valores contables, tales como una lista de enteros no negativos.

Con una distribución de probabilidad discreta, cada valor posible de la variable aleatoria discreta puede estar asociado con una probabilidad distinta de cero. Por lo tanto, una distribución de probabilidad discreta suele representarse en forma tabular.

Dentro de las distribuciones de variable discretas más importante están:

1) Distribución binomial. Mide el número de éxitos en una secuencia de n ensayos independientes de Bernoulli con una probabilidad fija p de ocurrencia del éxito entre los ensayos.

2) Distribución binomial negativa. Es una distribución de probabilidad discreta que incluye a la distribución de Pascal.

3) Distribución de poisson. Nos da a conocer, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo

4) Distribución geométrica. Es cualquiera de las dos distribuciones de probabilidad siguientes:

- La distribución de probabilidad del número X del ensayo de Bernoulli necesaria para obtener un éxito, contenido en el conjunto { 1, 2, 3,…} o

- La distribución de probabilidad del número Y = X − 1 de fallos antes del primer éxito, contenido en el conjunto { 0, 1, 2, 3,… }.

5) Distribución hipergeométrica. Es una distribución discreta relacionada con muestreos aleatorios y sin reemplazo.

6) Distribución uniforme discreta. Nombrada en honor al matemático y científico suizo Jakob Bernoulli, es una distribución de probabilidad discreta, que toma valor 1 para la probabilidad de éxito (p) y valor 0 para la probabilidad de fracaso (q = 1 − p).

Si X es una variable aleatoria que mide "número de éxitos", y se lleva a cabo un único experimento con dos posibles resultados (éxito o fracaso), se dice que la variable aleatoria X se distribuye como una Bernoulli de parámetro p.

Otro factor muy importante dentro de la estadística es el intervalo de confianza. Este se define como aquel que "describe la variabilidad entre la medida obtenida en un estudio y la medida real de la población (el valor real). Corresponde a un rango de valores, cuya distribución es normal y en el cual se encuentra, con alta probabilidad, el valor real de una determinada variable. Esta «alta probabilidad» se ha establecido por consenso en 95%.

Así, un intervalo de confianza de 95% nos indica que dentro del rango dado se encuentra el valor real de un parámetro con 95% de certeza"

Todas estas herramientas han sido creadas para facilitarnos la vida y en cualquier momento puede sacarnos de un apuro. Es fundamental saber aplicar las técnicas pues nos pueden llevar a evitar conflictos próximos en nuestra vida. Es vital que nosotros sepamos hacer cualquiera de las operaciones pues se nos pueden presentar diversas situaciones.

Fuentes consultadas:

https://plataformaeducativauds.com.mx/assets/biblioteca/6abce34f323672eed335f9f090e71fcd.pdf

https://support.minitab.com/es-mx/minitab/18/help-and-how-to/probability-distributions-and-random-data/supporting-topics/basics/continuous-and-discrete-probability-distributions/#:~:text=libras%20es%20cero.-,%C2%BFQu%C3%A9%20es%20una%20distribuci%C3%B3n%20discreta%3F,lista%20de%20enteros%20no%20negativos.

https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0034-98872005000900017

II. Resuelve los siguientes ejercicios.

1.- En la universidad del sureste hay 120 alumnos de bachillerato elegir una muestra de 20 alumnos para hacerles una serie de preguntas, utiliza el muestreo aleatorio simple y el muestreo aleatorio sistemático.

El primer método que emplearemos es el método simple.

Para ello los datos son:

N (población)= 120 (alumnos).

n (muestra)= 20.

120Ran#= número aleatorio.

120Ran#= 12.35= 12

120Ran#= 35.58= 36

120Ran#= 55.75= 56

120Ran#= 66.45= 66

120Ran#= 4.55= 5

120Ran#= 86.75= 87

120Ran#= 45.67= 46

120Ran#= 67.88= 68

120Ran#= 79.99= 80

120Ran#= 57. 34= 57

120Ran#= 111.12= 111

120Ran#= 9.99= 10

120Ran#= 38.34= 38

120Ran#= 34.45= 34

120Ran#= 21.59= 22

120Ran#= 102.89= 103

120Ran#= 12.94= 13

120Ran#= 29.66= 30

120Ran#= 89.56= 90

120Ran#= 77.45= 77

Muestra obtenida=

12, 36, 56, 66, 5, 87, 46, 68, 80, 57, 111, 10, 38, 34, 22, 103, 13, 30, 90, 77.

Sistemático:

K= N+K

K= 120/20 =6

Muestra obtenida=

6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90, 96, 102, 108, 114, 120.

2.- Si N=45 obtener una muestra de n=7 mediante el muestreo aleatorio simple.

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44.

45Ran#= 23.45= 23

45Ran#= 41.78= 41

45Ran#= 1.65= 2

45Ran#= 39.69= 38

45Ran#= 27.34= 27

45Ran#= 42. 23= 43

45Ran#= 17.78= 18

Muestra seleccionada= 23, 41, 2, 38, 27, 43, 18.

3.- En el evento estatal de fútbol interescolar, donde hubo una participación de 300 estudiantes, se registraron diversos incidentes, relacionados con diferentes tipos de lesiones físicas: 8 fracturas de diversas índoles, 35 dislocaciones, 40 torceduras, 26 tirones, 121 inflamaciones y los restantes participantes no sufrieron lesión alguna. Dados estos datos encuentra la distribución de probabilidad correspondiente.

Tipo de variable: Aleatoria discreta (tipos de lesiones).

Espacio muestral (posibles resultados a obtener) = fractura, dislocación, torcedura, tirón, inflamación y ninguna lesión.

Probabilidad de que el resultado sea fractura= P (x= fractura) = 8/300=0.026%

Probabilidad de que el resultado sea dislocación= P (x=dislocación) = 35/800=0.11%

Probabilidad de que el resultado sea torcedura= P (x=torcedura) = 40/300= 0.13%

Probabilidad de que el resultado sea tirón= P (x=tirón) = 26/300=0.086%

Probabilidad de que el resultado sea inflamación= P (x=inflación) = 121/300=0.40%

Probabilidad de que el resultado sea sin lesión= P (x=sin lesión) =70/300=0.23%

Nombre de la lesión	Probabilidad	Porcentaje de probabilidad
Fractura	8/300	0.026%
Dislocación	35/800	0.11%
Torcedura	40/800	0.13%
Tirón	26/800	0.086%
Inflamación	121/800	0.40%
Lesión	70/800	0.23%